1. Problem Statement:

The global pandemic has impacted all industries, with the hospitality sector experiencing some of the most significant challenges. As the world transitions into a post-pandemic phase, changes both minor and major are emerging within this sector. Signs of this transformation are visible, including governments authorizing the resumption of both domestic and international air travel. This development encourages individuals who have remained isolated for a considerable period to consider once again traveling for leisure. Consequently, analyzing the hotel reservation data, forming a cognition for the hotel business, finding key indicators, and making suggestions on how to further promote the hotel reservation service and build a hotel potential customer identification model is critical.

2. Articulation of value:

A. Revitalization of the Hospitality Sector

Post the pandemic, a well-analyzed strategy can lead to an economic revival of the hotel industry, generating revenue and providing employment opportunities. By encouraging more individuals to travel for leisure, the project can help reignite the tourism sector, which is intrinsically linked with the hospitality industry.

B. Enhancing Customer Experience

Leveraging data analytics can help hotels offer personalized services to their customers, enhancing their experience and potentially leading to higher customer retention. Furthermore, Customers are more concerned about safety and hygiene. Analytics can help in identifying and implementing measures that address these concerns, thus fostering trust and confidence in the hotel's services.

C. Strategic Decision Making

The project will enable stakeholders to make informed decisions based on data-driven insights, which can potentially lead to improved profitability and operational efficiencies. By identifying key indicators and customer preferences, the hotels can align their services to market demands, potentially gaining a competitive edge in the industry.

D. Innovation and Adaptability

Insights gathered from data analytics can pave the way for innovative service offerings, catering to the emerging needs and preferences of the customers in the world. Additionally, The project can facilitate the

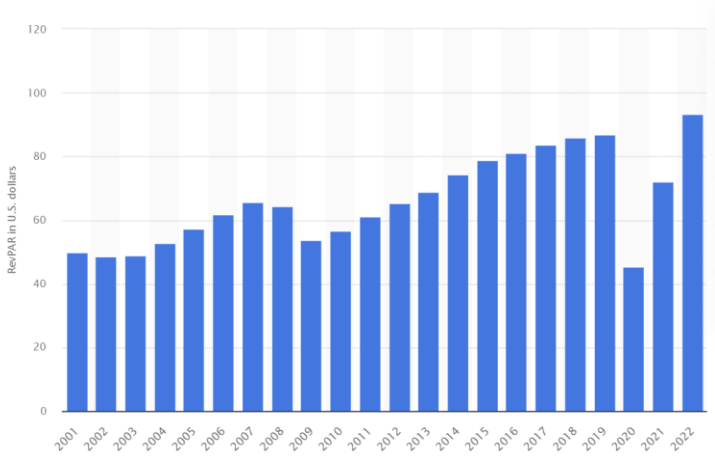
integration of modern technologies in the hotel industry, like AI-driven customer service, enhancing operational efficiency and customer satisfaction.

3. Calculation of the potential economic value:

A. Assumptions

a. Increase in Hotel Reservations: We assume a 20% year-over-year increase in hotel reservations owing to the resumption of travel activities.

b. Average Revenue Per Booking: Based on historical data, the average revenue per booking is assumed to be 93.27 U.S. dollars.[1]



c. Customer Retention Rate: Through personalized services and improved customer experience, we assume an increase in the customer retention rate by 55%.[2]

d. Operational Efficiency[3]: Implementing data-driven strategies and technological integration is assumed to improve operational efficiency, potentially reducing operational costs by 10%.

B. Calculations

a. Projected Revenue Increase

Projected Revenue Increase=Number of Additional Bookings*Average Revenue Per Booking

(Assuming the number of additional bookings to be 10,000, Projected Revenue Increase=10,000×93.27=$932,700)

b. Increased Revenue from Customer Retention

Revenue from Customer Retention=Retention Rate Increase*Current Customer Base*Average Revenue Per Booking

(Assuming the current customer base to be 50,000, Revenue from Customer Retention=0.55×50,000×93.27=$2,564,925)

c. Savings from Operational Efficiency

Operational Savings=Operational Cost Reduction*Current Operational Costs

(Assuming the current operational costs to be $5,000,000, Operational Savings=0.10×5,000,000=$500,000)

d. Total Potential Economic Value

The total potential economic value can be calculated by adding the projected revenue increase, the revenue from customer retention, and the operational savings.

(Total Potential Economic Value=$932,700+$2,564,925+$500,000=$3,997,625)

C. Footnotes

a. The growth rates and percentages used in the calculations are assumed based on the current market trends and should be verified with actual data for precision.

b. The average revenue per booking is a hypothetical figure and should be replaced with the real average value based on historical data.

c. The current operational costs are assumed for this calculation and should be replaced with the actual operational costs.

4. Project plan:

WEEK 1	Identify the problem statement and dataset
WEEK 2	Ingest and explore the dataset
WEEK 3	Perform exploratory data analysis
WEEK 4	Make data model ready
WEEK 5	Engineer features
WEEK 6	Develop 1st modeling approach (simple, the baseline)
WEEK 7	Develop 2nd modeling approach (more complex)
WEEK 8	Develop 3rd modeling approach (even more complex)
WEEK 9	Select the winning model
WEEK 10	Data Centric AI
WEEK 11	Explain the model, analyze risk, bias, and ethical considerations
WEEK 12	Save and package your model for deployment. Build your model monitoring plan
WEEK 13	Bring it all together

5. Discuss the dataset:

A. Dataset Overview: The dataset in discussion encapsulates detailed information pertaining to reservations made at a Resort Hotel. This comprehensive dataset includes a myriad of features that illuminate different aspects of the booking process, offering a rich foundation for multifaceted analysis in understanding and fostering growth in the hotel business in the post-pandemic phase. Here is a brief explanation of how some of the key features could be utilized to analyze and strategize:

a. Booking and Stay Details: Features such as lead_time, arrival_date_*, stays_in_weekend_nights, and stays_in_week_nights will facilitate the analysis of booking trends, helping in forecasting demands and strategizing pricing during peak periods.

b. Guest Details: Information like adults, children, babies, country, and contact details (name, email, and phone-number) can be leveraged to understand the demographics of the guests, which is critical in personalizing services and marketing efforts.

c. Booking Preferences and Behavior: Features like meal, reserved_room_type, assigned_room_type, and total_of_special_requests will provide insights into guest preferences, assisting in enhancing customer satisfaction and tailor services accordingly.

d. Market and Distribution Analysis: The market_segment and distribution_channel attributes can be used to analyze the effectiveness of different market strategies and distribution channels, helping to optimize marketing strategies and enhance outreach.

e. Customer Loyalty and Retention: Attributes like is_repeated_guest, previous_cancellations, and previous_bookings_not_canceled will offer a deeper understanding of customer loyalty and behaviors, vital in building customer retention strategies.

B. Data and Application:

a. Market Analysis & Strategy Formulation:

Seasonality Analysis: By analyzing attributes like arrival_date_month, arrival_date_week_number, etc., patterns of seasonality can be identified. This information can guide strategies for dynamic pricing and promotions.

Customer Segmentation: Utilizing details about the guests (e.g., country, market_segment) can help in identifying distinct customer segments, allowing for more targeted marketing strategies.

b. Building a Potential Customer Identification Model:

Data-Driven Insights: The dataset holds critical information to build a potential customer identification model. Attributes like is_repeated_guest, market_segment, distribution_channel, etc., can be used in predictive modeling to identify potential customers and craft strategies to attract them.

6. Identify the type of modeling:

A. Supervised Learning: Given that one of the objectives is to build a hotel potential customer identification model, I am looking at a classification problem within supervised learning.

a. Binary Classification: The attribute is_repeated_guest can be utilized in a binary classification model where we aim to predict whether a guest will be a repeat customer or not, based on various other features in the dataset.

b. Multi-Class Classification: I will develop a multi-class classification model to segment the guests into different categories based on their preferences and behaviors, which can be identified using attributes like market_segment, country, meal type, and more. This segmentation can help in crafting targeted marketing strategies.

B. Unsupervised Learning: Alongside this, I can employ unsupervised learning methods to uncover hidden patterns and relationships within the data that might not be apparent or well-defined yet. Here are some potential approaches:

a. Clustering: Clustering algorithms, like K-means, can be used to segment the customer base into distinct groups based on similarities in their booking behaviors and preferences. This can aid in identifying niche market segments and customizing offerings accordingly.

b. Association Rule Mining: This can help in identifying rules that highlight relationships between seemingly independent data in the dataset. For example, understanding the association between the reserved_room_type and total_of_special_requests to offer package deals that are more attractive to the customers.

c. Dimensionality Reduction: Techniques such as PCA (Principal Component Analysis) can be used to reduce the dimensionality of the data while retaining most of the important information. This can be particularly useful in visualizing high-dimensional data and making the models more efficient.

Reference:

[1] Revenue per available room in the U.S. hotel industry 2001-2022 (Published by Statista Research Department, Jun 15, 2023).

[2] https://info.liftoff.io/2020-mobile-app-trends-report.

[3] https://www.hospitalitynet.org/news/4109214.html.